Whether “Clever Hans” effect affect fitting a logistic regression model to the dataset ex2120

Chiu Fan Hui

13/10/2020

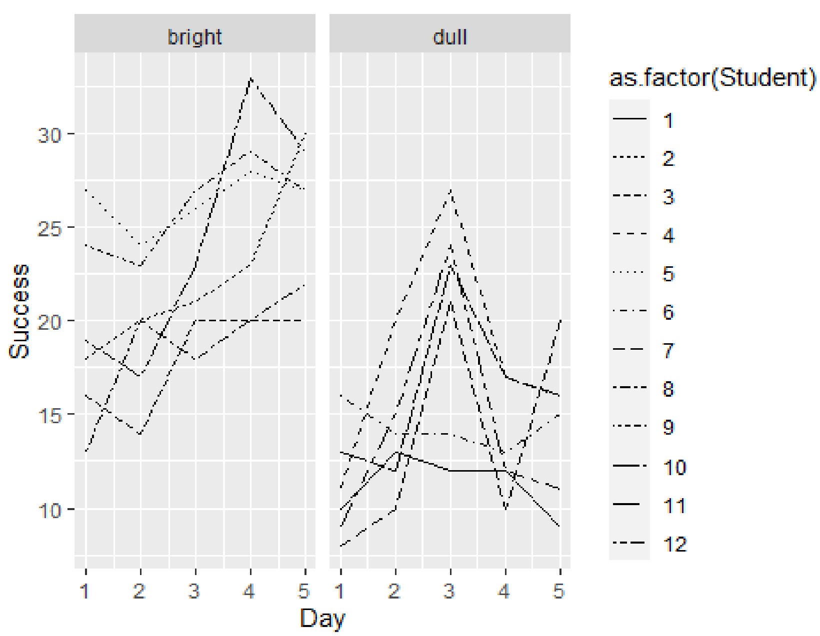
Aim: Investigating whether there is evidence for a “Clever Hans” effect for fitting a logistic regression model to the dataset ex2120

Background: Data set ex2120 is used in this project, each of 12 students trained rats to run a maze, the data set contains their number of successful runs out of 50 on each of 5 days, the student’s prior expectation of success (on a scale from -10 to 10), and a variable indicating treatment–whether or not the students were supplied with the fictitious information that their rights were bright.

Let’s explore the data first

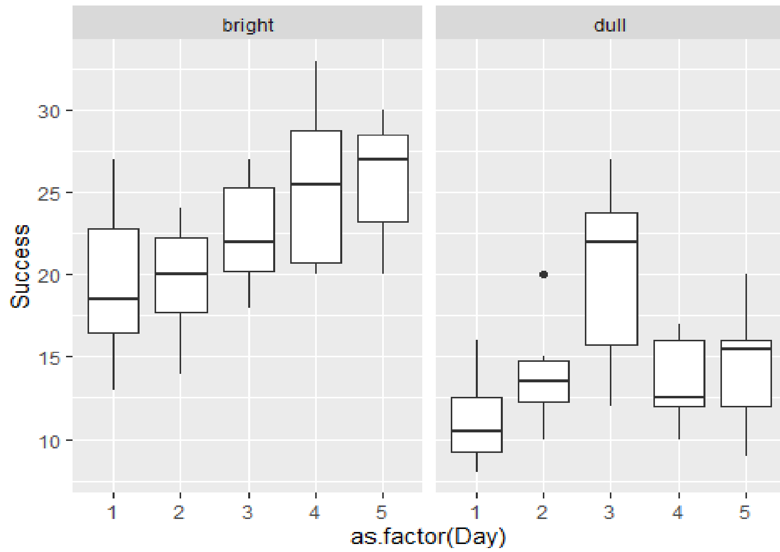
```
library(Sleuth3)
library(ggplot2)
```

```
#Visualising the data
p <- ggplot(ex2120, aes(x=Day, y=Success, linetype=as.factor(Student)))
p + geom_line() + facet_grid(~Treatment)
```



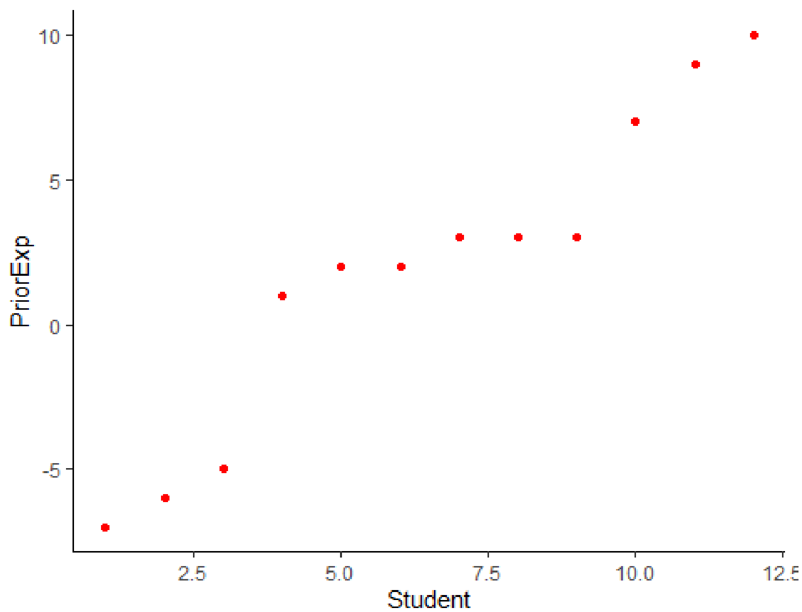
In average, the number of successes for bright rats is higher than dull rats. Most of the dull rats had a peak performance on day three.

```
p <- ggplot(ex2120, aes(x=as.factor(Day), y=Success))
p + geom_boxplot() + facet_grid(~Treatment)
```



The dull rats were doing well on day three, as the box is significantly higher than the rest of the day. The median of bright rats seems increasing, but most of the box is overlapping. Bright rats had a better performance on day five compare to day one and two as the box of day five is significantly higher than day one and two.

```
#why we don't want to use Student as a variable
p <- ggplot(ex2120, aes(x=Student, y=PriorExp))
p + geom_point(colour="red") + theme_classic()
```



Each student has a different prior expectation, so the prior expectation will be used for a model, not the student.

```r
#1. Using the week 10 notes
tapply(ex2120$Success, ex2120$Treatment, sum)

## bright   dull
##    678    440

odds_bright = 678/(1500 - 678)
odds_dull = 440/(1500 - 440)
or <- odds_bright/odds_dull
lor <- log(or)
se.lor <- sqrt(1/(678*822/1500) + 1/(440*1060/1500))
approx.95CI <- c(lor + qnorm(0.025)*se.lor, lor + qnorm(0.975)*se.lor)
or

## [1] 1.98706

exp(approx.95CI)

## [1] 1.709177 2.310123
```

The total number of successes for the bright rats are 678, for the dull rats are 440. The odds ratio is 1.98706, not equal to one, which mean the odds for bright rats and the odds for dull rats are different. As the exponetial of the confidence interval (1.709177, 2.310123) does not include one, so the odds for bright rats is not equal to the odds for dull rats.

```r
#2. Using fisher.test
rats <- matrix(c(678, 1500-678, 440, 1500-440), nrow = 2)
rats
```

```
##      [,1] [,2]
## [1,]  678  440
## [2,]  822 1060
```

```r
fisher.test(rats) #p valuse is very low and true odds ratio is not equa
l to 1 which mean odds for bright not equal odds for dull
```

```
## 
##  Fisher's Exact Test for Count Data
## 
## data:  rats
## p-value < 2.2e-16
## alternative hypothesis: true odds ratio is not equal to 1
## 95 percent confidence interval:
##  1.704414 2.317073
## sample estimates:
## odds ratio 
##   1.986617
```

There are 678 bright rats succeed and 822 bright rats failed. There are 440 dull rats succeed and 1060 dull rats failed. The p-value is < 2.2e-16 so the null hypothesis can be rejected and the true odds ratio is not equal to 1, which mean the odds for bright rats and the odds for dull rats are different.

```r
#3. via glm
binResponse <- cbind(ex2120$Success, 50 - ex2120$Success)
ex2120$Treatment <- factor(ex2120$Treatment, levels=c("dull", "bright")
)
fit <- glm(binResponse ~ Treatment, family = binomial(link=logit), data
 = ex2120)
exp(confint(fit))
```

```
##                    2.5 %    97.5 %
## (Intercept)     0.371100 0.4635131
## Treatmentbright 1.709803 2.3110733
```

```r
summary(fit)
```

```
##
## Call:
## glm(formula = binResponse ~ Treatment, family = binomial(link = logi
t),
##     data = ex2120)
##
## Deviance Residuals:
##    Min      1Q  Median      3Q     Max
## -2.802  -0.891  -0.208   1.034   3.634
##
## Coefficients:
##                 Estimate Std. Error z value Pr(>|z|)
## (Intercept)     -0.87925    0.05671 -15.504   <2e-16 ***
## Treatmentbright  0.68666    0.07686   8.934   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 201.54  on 59  degrees of freedom
## Residual deviance: 120.31  on 58  degrees of freedom
## AIC: 378.02
##
## Number of Fisher Scoring iterations: 4
```

As the exponetial of the confidence interval (1.709177, 2.310123) does not include one, so the odds for bright rats is not equal to the odds for dull rats. Both the Pr(>|z|) are <2e-16, so the null hypothesis can be rejected and the true odds ratio is not equal to 1, which mean the odds for bright rats and the odds for dull rats are different.

```r
#construct a model
fit <- glm(binResponse ~ Treatment + Day + PriorExp, family=binomial(li
nk=logit), data=ex2120)
summary(fit)
```

```
##
## Call:
## glm(formula = binResponse ~ Treatment + Day + PriorExp, family = bin
omial(link = logit),
##     data = ex2120)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -2.2803  -0.8378  -0.1295   0.7614   3.7635
##
## Coefficients:
##                 Estimate Std. Error z value Pr(>|z|)
## (Intercept)     -1.26553    0.10370 -12.204  < 2e-16 ***
## Treatmentbright  0.71976    0.07798   9.230  < 2e-16 ***
## Day              0.10910    0.02725   4.004 6.22e-05 ***
## PriorExp         0.02071    0.00736   2.814   0.0049 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 201.543  on 59  degrees of freedom
## Residual deviance:  96.264  on 56  degrees of freedom
## AIC: 357.98
##
## Number of Fisher Scoring iterations: 4
```

```r
drop1(fit, test="Chi")
```

```
## Single term deletions
##
## Model:
## binResponse ~ Treatment + Day + PriorExp
##           Df Deviance    AIC    LRT  Pr(>Chi)
## <none>         96.264 357.98
## Treatment  1  183.250 442.96 86.986 < 2.2e-16 ***
## Day        1  112.383 372.10 16.119  5.95e-05 ***
## PriorExp   1  104.234 363.95  7.970  0.004756 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
residual.deviance <- summary(fit)$deviance
deg.of.freedom <- summary(fit)$df.residual
#pvalue
1 - pchisq(residual.deviance,deg.of.freedom)
```

```
## [1] 0.0006624085
#no it doesn't pass
```

The AIC is 357.98, and the residual deviance is 96.264 on 56 degrees of freedom. The p-value is 0.0006624085, it's highly significant, which means there are a large amount of residual deviance that is not explain in this model. Therefore, this model fail the goodness of fit test.

```
#Make Day a factor
fit.dayfac <- glm(binResponse ~ Treatment + as.factor(Day) + PriorExp,
family=binomial(link=logit), data=ex2120)
summary(fit.dayfac)

##
## Call:
## glm(formula = binResponse ~ Treatment + as.factor(Day) + PriorExp,
##     family = binomial(link = logit), data = ex2120)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -2.2750  -0.9184   0.0360   0.6931   2.9640
##
## Coefficients:
##                   Estimate Std. Error z value Pr(>|z|)
## (Intercept)      -1.240508   0.102033 -12.158  < 2e-16 ***
## Treatmentbright   0.722176   0.078120   9.244  < 2e-16 ***
## as.factor(Day)2   0.141523   0.125468   1.128 0.259335
## as.factor(Day)3   0.536164   0.122899   4.363 1.28e-05 ***
## as.factor(Day)4   0.379453   0.123654   3.069 0.002150 **
## as.factor(Day)5   0.436938   0.123338   3.543 0.000396 ***
## PriorExp          0.020776   0.007372   2.818 0.004829 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 201.54  on 59  degrees of freedom
## Residual deviance:  86.23  on 53  degrees of freedom
## AIC: 353.94
##
## Number of Fisher Scoring iterations: 4

drop1(fit.dayfac, test="Chi")

## Single term deletions
##
## Model:
## binResponse ~ Treatment + as.factor(Day) + PriorExp
##                  Df Deviance    AIC    LRT  Pr(>Chi)
## <none>              86.230 353.94
```

```
## Treatment       1  173.502 439.21 87.272 < 2.2e-16 ***
## as.factor(Day)  4  112.383 372.10 26.152 2.948e-05 ***
## PriorExp        1   94.227 359.94  7.997  0.004686 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

#recheck GOF
residual.deviance <- summary(fit.dayfac)$deviance
deg.of.freedom <- summary(fit.dayfac)$df.residual
#pvalue
1 - pchisq(residual.deviance,deg.of.freedom)

## [1] 0.00264833

#better but it still doesn't pass
```

The AIC is 353.94, the residual deviance is 86.23 on 53 degrees of freedom and the p-value is 0.00264833.

When the Day is set as a factor, AIC will be lower, residual deviance is less and p-value is better, so the model do a little bit better. Interaction are tried in this project but they are not significant

```r
#look at the CIs
exp(confint(fit)) #if not contain 1 then it's highly sinicicant

##                      2.5 %    97.5 %
## (Intercept)     0.2298597 0.3451752
## Treatmentbright 1.7635761 2.3942265
## Day             1.0573645 1.1765733
## PriorExp        1.0063375 1.0358002

#can we plot the predictions of the odel
priorexp <- seq(-7,10,1)
day <- seq(1,5,1)
treat <- c("bright","dull")
grid <- expand.grid(PriorExp=priorexp, Treatment=treat, Day=day)
```

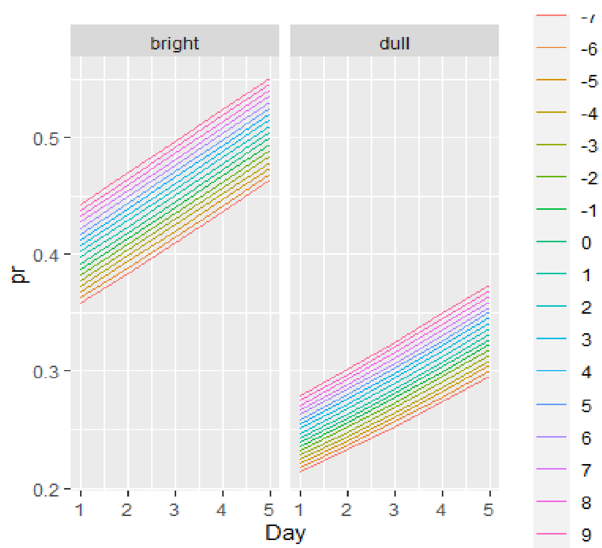
The confidence interval do not include one so they are all significant.

```r
pr<- predict(fit, newdata=grid, type="response")

toPlot <- cbind(grid, pr)
head(toPlot)

##   PriorExp Treatment Day        pr
## 1       -7    bright   1 0.3585597
## 2       -6    bright   1 0.3633359
## 3       -5    bright   1 0.3681392
## 4       -4    bright   1 0.3729688
## 5       -3    bright   1 0.3778239
## 6       -2    bright   1 0.3827036

p <- ggplot(toPlot, aes(x=Day, y=pr, color=as.factor(PriorExp)))
p + geom_line() + facet_grid(~Treatment)
```

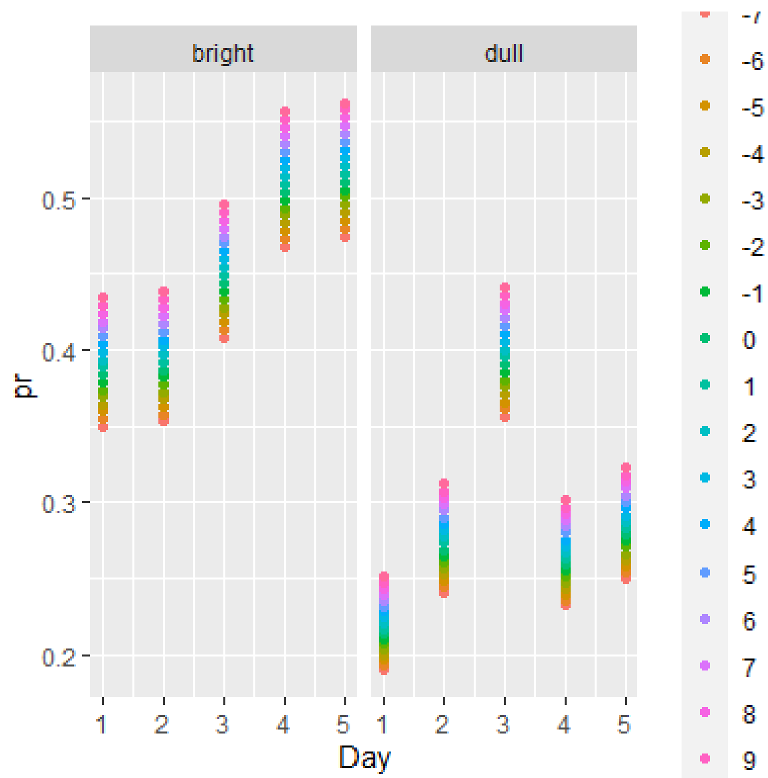


The prediction of the proportion of success show they all increase, but dull rats have a lower mean than bright rats.

```r
#plot the predictions of he day as factor + interaction w Treatment mod
el
fit.dayfac2 <- glm(binResponse ~ Treatment + as.factor(Day) + PriorExp
+ Treatment:as.factor(Day), family=binomial(link=logit), data=ex2120)
pr<- predict(fit.dayfac2, newdata=grid, type="response")
toPlot <- cbind(grid, pr)
head(toPlot)
```

```
##   PriorExp Treatment Day        pr
## 1       -7    bright   1 0.3499422
## 2       -6    bright   1 0.3547089
## 3       -5    bright   1 0.3595046
## 4       -4    bright   1 0.3643285
## 5       -3    bright   1 0.3691799
## 6       -2    bright   1 0.3740578
```

```r
p <- ggplot(toPlot, aes(x=Day, y=pr, color=as.factor(PriorExp)))
p + geom_point() + facet_grid(~Treatment)
```



When Day is set as a factor, there is a similar trend as before

```r
# Visualising predictions
pred.response <- predict(fit, type="response")
data.frame(prediction=pred.response, status=ex2120$Success/50)
```

```
##    prediction status
## 1   0.2139332   0.20
## 2   0.2328519   0.26
## 3   0.2529055   0.24
## 4   0.2740690   0.24
## 5   0.2963012   0.18
## 6   0.3633359   0.36
## 7   0.3889297   0.40
## 8   0.4151508   0.42
## 9   0.4418614   0.46
## 10  0.4689127   0.60
## 11  0.3681392   0.32
## 12  0.3938621   0.28
## 13  0.4201870   0.40
## 14  0.4469740   0.40
## 15  0.4740724   0.40
## 16  0.2431060   0.22
## 17  0.2637390   0.40
## 18  0.2854628   0.54
## 19  0.3082268   0.34
## 20  0.3319627   0.32
## 21  0.4024512   0.54
## 22  0.4289440   0.48
## 23  0.4558503   0.52
## 24  0.4830169   0.56
## 25  0.5102842   0.54
## 26  0.2469363   0.32
## 27  0.2677794   0.28
## 28  0.2897050   0.28
## 29  0.3126593   0.26
## 30  0.3365705   0.30
## 31  0.2508070   0.16
## 32  0.2718589   0.20
## 33  0.2939844   0.42
## 34  0.3171264   0.20
## 35  0.3412097   0.40
## 36  0.4074407   0.26
## 37  0.4340234   0.40
## 38  0.4609911   0.36
## 39  0.4881892   0.40
## 40  0.5154573   0.44
## 41  0.4074407   0.48
## 42  0.4340234   0.46
## 43  0.4609911   0.54
## 44  0.4881892   0.58
## 45  0.5154573   0.54
```

```
## 46  0.2666890   0.26
## 47  0.2885605   0.24
## 48  0.3114639   0.46
## 49  0.3353283   0.34
## 50  0.3600650   0.32
## 51  0.2748659   0.18
## 52  0.2971362   0.30
## 53  0.3204137   0.48
## 54  0.3446206   0.24
## 55  0.3696616   0.22
## 56  0.4428484   0.38
## 57  0.4699093   0.34
## 58  0.4971482   0.46
## 59  0.5244040   0.66
## 60  0.5515152   0.58
```

The prediction are similar to the status, the modle make a good prediction